

“New” Rankings on the Scene: The U21 Ranking of National Higher Education Systems and U-Multirank

Gergely Kováts

1 Introduction

Although the ranking of higher education institutions (HEIs) has a history of several decades in some countries (Salmi and Saroyan 2007), the global public awareness has been increasing since the appearance of the first truly global ranking, the Academic Ranking of World Universities (ARWU) in 2003. Since then there has been an explosion of national and international rankings combined with growing international interest about this phenomenon. A good example of it being foundation of International Ranking Expert Group (2002) and the formulation of the so-called Berlin Principles (2006).

But why do rankings become so popular in such a short time? One possible reason can be the rising information demand regarding institutions and the changing role of prestige in higher education.

The expansion of higher education resulted in an increasing number and diversity of students, institutions and study programmes leading to an increased complexity of the sector. This is especially true for Europe (and the European Higher Education Area), where each national higher education system evolved in a more or less unique way. Complexity is exacerbated by information asymmetry, because higher education provides so-called “experience goods”, that is, one can evaluate (partially) the service of an institution if he/she tries it out. Once admitted to an institution, however, it is not easy to change to another one for example because of the sunk costs (even if initiatives such as the credit system aims to reduce this lock-in effect), as these costs increase the requirements of information before applying to an institution. Globalization and diminishing borders have made higher education institutions abroad available for mobile students, thus increasing information needs even further.

G. Kováts (✉)

Corvinus University of Budapest, Budapest, Hungary
e-mail: gergely.kovats@uni-corvinus.hu

For employers, the expansion of higher education has made the evaluation of the quality of graduated students and institutional research performance more important. In addition, the needs of increased public and private funding draw governments' attention to the transparency and accountability of higher education institutions.

In sum, on one hand, there is a growing demand for information and transparency. On the other hand, however, the performance and quality of higher education institutions are more and more difficult to assess in a complex environment.

As a result, many transparency tools have been developed, such as recognition, ECTS credits system, qualifications frameworks, learning outcomes, three cycles system, diploma supplement, national admission websites, higher education institutions' websites, study guides, annual reports (Hazelkorn 2012; Vercruyssen and Proteasa 2012). External assessment tools and procedures are part of transparency tools: in addition to program and institutional accreditations, audits of QA systems, reporting practices, regular statistical data provisions and financial monitoring etc., there are rankings, which offer simple and convenient methods to grab the essence of institutional performance and quality. Moreover, they also allow the comparison of institutions.

From the institutional point of view, however, the importance of rankings can be explained a bit differently. Because of performance assessment problems, the differentiation of institutions is difficult. It is not just the performance and quality of an institution that matters when one has to choose between institutions, but also the *appearance* of performance and quality, which is reflected in the prestige of an institution.

The prestige of a higher education institution expresses to what extent the organization meets and surpasses the expectations regarding higher education as a social institution. These expectations embody what higher education institutions should do and how they should do it, therefore they define the standards, as well as the frame of reference of institutional performance and quality. As prestige, which includes legitimacy, status and reputation of an organization, determines "an organization's capacity to achieve objectives by virtue of enjoying a favourable social evaluation" (Deephouse and Suchman 2006). Prestige has a huge impact on the capability of an organization to attract further (state or third-party) resources which can be used to enhance further its prestige, leading to a continuous "reputation race" (van Vught 2008) and to the emergence of a winner-take-all market (Eckel 2008).

2 The General Characteristics of Rankings

Current rankings have been playing a pivotal role in creating and conferring prestige to institutions. To sum up, demand from stakeholders, as well as from institutions, results in the proliferation of rankings.

Their impact could be illustrated by many examples (Hazelkorn 2011; Rauhvargers 2013; Sadlak 2014; Salmi and Saroyan 2007):

- institutional management treats the improvement in major rankings as a strategic goal, establishes offices to collect data and to track progress,
- boards bind bonuses or further employment of senior managers to improvement in rankings,
- policy initiatives, such as the Excellent Initiatives in Germany, Project 985 in China¹ or mergers (e.g. Aalto University or University of Manchester), explicitly aim to increase the number of world-class (that is, better ranked) universities or improving ranking position,
- immigration regulations and state scholarship programmes increasingly take into consideration the international ranking of institutions to determine the quality of institutions etc.

Many types of rankings exist. There are national, regional and international rankings. Rankings may focus on a special group of institutions (e.g. business schools, young institutions) or the higher education sector as a whole. Some of them rank institutions, while others rank faculties or educational programmes. Ranking may focus on one aspect of institutional activity (research mostly), or many different facets at the same time.

Although the number and type of rankings may be high, they have some common characteristics. Rankings

- are summative in nature (rather than formative), that is, they judge institutions by their past performance,
- focus on comparing entities (rather than enhancing and improving them),
- are produced by external assessors, even if institutional cooperation is required (e.g. data provision) and
- institutions in rankings are identifiable (not anonymous).

The most prominent global, institutional rankings—such as the ARWU, Quacquarelli Symonds’s ranking (QS) and the ranking of the Times Higher Education (THE)—and several other (national) ones share the following additional characteristics:

- they are public, rather than open for only a narrow audience (e.g. the government, institutions themselves etc.),
- they are hierarchical, as they want to order institutions (rather than rate or categorize them, for example),
- they produce one overall ranking, even if they use many different indicators to grab different facets of institutional activities,
- they are competitive, that is, there is only one No. 1 institution,
- participation is voluntary (not obligatory) or does not require institutional cooperation.

¹See: http://en.wikipedia.org/wiki/Project_985 (accessed: 25 Sept, 2014).

It is worth noting however, that most global ranking providers have more than one product. For example, QS has a non-competitive, external assessment service where institutions may earn stars (QS stars) based on their assessment etc. (for an overview, see Rauhvargers 2013). The main products, which draw most of the attention, are still global institutional rankings.

3 The Criticism of Rankings

Despite the growing demands, the proliferation of rankings and the beneficial effects (such as more conscious strategic management, development of reporting and data gathering procedures, dialogue about quality and performance, consumer guidance), criticism is also widespread. There are conceptual and methodological concerns.

One of the conceptual problems is that current rankings strengthen hierarchical stratification instead of acknowledging horizontal diversity (van Vught and Ziegele 2011). Rankings do not simply provide an overview of performance and quality according to current standards and expectations, they also create, shape and legitimize those expectations. Ranking providers have recently emphasised that they focus only on global research institutions rather than all institutions. However, the choice of names (e.g. “world university ranking” is usually included in their names) suggest differently, and in public discourses these global rankings are usually interpreted as rankings of institutions that matter (in general). As a result, institutions face expectations tailored to the international research universities, as most indicators favour this type of institution (e.g. indicators regarding internationalization, the amount and impact of research, number of academics with Nobel Prize etc.). Rankings, therefore, make international research universities a “single global status model” (van der Wende 2008) for everyone, suggesting there is only one way to be a good institution: to imitate the No. 1. university. As Hazelkorn wrote “institutions are essentially ranked according to how much they deviate from the ‘best’; in other words, to what extent are universities at variance with Harvard?” (Hazelkorn 2011).

By implicitly setting global standards, rankings also contribute to the social construction of reputation race and winner-take-all-market, that is, to an increasing vertical stratification, where few highly prestigious (‘world-class’) institutions emerge and steadily increase their advantage, while others drop behind despite their efforts. It is easy to hypothesize that “academic drift” becomes more intense, because institutions with profiles different from the global model are forced to become similar to it or else they get stuck in a disadvantageous position. In both cases, the result is the weakening of the diversity of higher education.

In addition, global rankings are insensitive to contextual differences. In some countries research is concentrated on universities, in other countries it is divided between universities and a network of independent research institutions. Institutions

funded mostly by the state and institutions from developing countries (where funding of research is scarce) are also adversely effected.

Hazelkorn (2011) emphasises that current rankings favour the traditional, Mode 1 Knowledge Production (Gibbons et al. 1994) because the results of this type of research are manifested in articles and books (which can be easily counted). The output of Mode 2 Knowledge Production, where problems are defined in the “context of application”, is the impact which is not necessarily generalized and published.²

Another conceptual problem is the search of “the best” university. In this endeavour current rankings produce an overall ranking by weighting indicators and creating a composite indicator. Different stakeholders, however, define “best” differently, and overall rankings make it impossible to enforce these differences. There are other methodological problems with composite indicators. The selection of weightings is arbitrary and depends solely on the preferences of the producer (Harvey 2008; van Vught and Ziegele 2011). Composite indicators also suggest the possibility of compensation, that is, bad results in one indicator can be counterbalanced by good performance in others. As a result, institutions with similar rank may have highly different profiles. Finally, the correlation between weighted indicators is usually strong, therefore some activities are taken into consideration more than once (Soh 2011).

Other frequently mentioned methodological problems are the following (Harvey 2008; Hazelkorn 2011; Rauhvargers 2011, 2013; van Vught and Ziegele 2011):

- the selection of indicators depends on what is measurable, and less on what is important. Important factors (such as indicators on teaching and learning experience) are omitted or included through proxy variable which causes distortion (e.g. measuring teaching quality with amount of resources per student or with student/staff ratio).
- bias of language/discipline: measuring research output in social sciences, humanities and arts is more difficult, because in these disciplines the role of book and book chapters is more important, but databases regarding these types of publications are incomplete. Therefore only those rankings are fair, where institutions with similar disciplines are compared. Another problem is the dominance of English language in research and in the international publication databases which affects favourably those countries where the native language is English. Rankings do not reflect on this distortion adequately (Rauhvargers 2013).
- data collection problems: some rankings (such as QS and THE) use the results of reputational surveys distributed among academics and employers. Low response rate, geographical dispersion of responses and halo effect gives rise to worry: the current prestige of an institution influences responses independently

²The Research Excellence Framework (REF), which is a regular rating procedure of institutional research productivity and capacity in the UK, has an interesting attempt to capture the Mode 2 research in some disciplines by requesting “impact case studies”. Results of the predecessor of REF (called Research Assessment Exercise) were frequently included in UK national rankings.

of its real performance (there are anecdotal cases where institutions ranked highly on those fields where they do not offer teaching programs or pursuit research; Hazelkorn 2011; Rauhvargers 2011). It is also questionable whether an academic can truly assess whole institutions (van der Wende 2008). Regarding reputational rankings it is especially true that they not just simply measure, but also reinforce current status quo (Rauhvargers 2013).

- consistency of institutional data: some rankings require institutional data provision. The condition of valid and reliable comparison is the consistency of data, which is hard to maintain if the number of international participant is high. In addition to the intentional data manipulation attempts, the lack of shared and mutual understanding of required data threatens consistency.
- frequently changing methodology is a problem in rankings which use composite indicators, because trends might be misleading. Changes may stem from efforts to improve methodology, but if the ranking position of an institution changes, it is hard to separate the effect of changing methodology from the effect of institutional responses. Hazelkorn even raises the possibility that ranking providers sometime change methodology intentionally to create news about changing ranking position (Hazelkorn 2011; van Vught and Ziegele 2011).
- the problem of distances: rankings indicate statistically non-significant differences as real. Distorting distances also hide vertical stratification, because they hide the distance between different ranking positions.
- lack of clarity: transparency of methodology, handling of missing data (Harvey 2008), the selection of ranked institution, as well as the eligibility criteria to be included in the ranking are rarely described clearly.

Based on these criticisms, however, new rankings have been developed in the last couple of years: U21 Ranking of National Higher Education Systems (U21) and U-Multirank. In the following sections it will be explored in more details to what extent these new rankings are able to overcome the problems of previous rankings and what strengths and shortcomings they have on methodological and conceptual level.

4 U21 Ranking of National Higher Education Systems

Traditional rankings focus on institutions. It is a frequent mistake to project the results of these rankings on national higher education systems drawing the false conclusion that a country has a world class higher education system if it has world class universities. Bad results on ARWU and other global rankings inspired many politicians to intervene. They launch excellence programs and mergers in order to improve ranking position, creating tensions within their higher education system as other institutions feel to be neglected (cf. Aula and Tienari 2011 summarized reactions on the foundation and increased state support of Aalto University in Finland). This is the result of a steeper vertical stratification.

Robert Birnbaum also draws attention to the misunderstood relationship between world class universities and higher education systems when he states that “the United States doesn’t have a world-class higher education system because it has many world-class universities; instead it has world-class universities because it has a world-class higher education system.” (Hazelkorn 2011).

The misunderstanding stems from the lack of focus of current rankings. Looking for “the best” universities, they become insensitive for the different demands of different stakeholders. Lack of clarification deludes governments who should focus on the world class systems rather than world class universities. This shortcoming is addressed by U21, a global network of research-intensive universities, which first sponsored the ranking of national higher education systems in 2012.

4.1 General Characteristics

In the 2014 report (Williams et al. 2014) 50 countries are ranked by weighting 24 indicators in 4 dimensions. The dimension of “resources” (weight of 20 %) with 5 indicators represents expenditure on higher education or on research and development in relative terms (per capita bases or as a percentage of GDP). The dimension of “environment” (weight of 20 %) has 4 indicators, the most interesting of which is the qualitative measure of the policy and regulatory environment, which mostly refers to the diversity and autonomy of institutions. There are two indicators for proportion of female student and female academics, and one for “data quality”. “Connectivity” (20 %) is a dimension with 6 indicators standing for the proportion of international students, number of co-authored publications with international collaborators or industry researchers, presence of institutions on the web and the rating of business executives regarding knowledge transfer between industry and universities. Finally, the dimension of “output” (40 %) has 9 indicators focusing on research performance and excellence, number of students and researchers, rate of graduate unemployment.

In the report of 2014, tables were also presented in which levels of economic development were taken into consideration. These tables show whether a country performs better or worse than is expected at their level of GDP. This addition makes the ranking more insightful.

The source of data for the majority of indicators is the database of one of the major international organizations (e.g. OECD, World Bank, UNESCO, ILO etc.) which is not just cost-efficient, as these data do not require additional efforts to collect, but with the exception of some cases,³ it also guarantees high degree of validity and consistency. Results of other rankings, such as ARWU, SCImago, Webometrics and Leiden

³Results taken from The Global Competitiveness Report (World Economic Forum) represents the opinion of business executives. It is a question whether these opinions are comparable on international level (Rauhvargers 2013).

Ranking are also incorporated. The indicator of “policy and regulatory environment” is calculated in a qualitative way by using expert opinions.

4.2 *Evaluation of U21*

The overall ranking of countries is the result of weighting of indicators. It is not surprising therefore that U21 faces similar methodological problems as those conventional institutional rankings which use composite indicators. Most notably,

- the selection of weightings is arbitrary.
- the correlation of indicators: Soh (2012) points at the fact that there is an underlying input-output model behind U21, where resources, environment and connectivity result in output. Thus, in the overall ranking output is counted twice: directly and indirectly.
- the selection of indicators is quite innovative, but it is guided by availability. For example teaching, teaching quality and learning are completely omitted from the ranking because there is no reliable international survey or ranking dealing with them. On the other hand, U21 might encourage countries and international data providers to collect more profound data. One reason for having only 50 countries included in the U21 ranking is the lack of data for the rest of the countries, which shed light on the quality of data in less developed countries.
- the methodology of U21 changed in every year: new indicators were introduced, weightings of dimensions and the handling of missing values were modified. The ranking position of few countries (Thailand, Taiwan-China) fluctuated, while others lost/gained considerable positions. For example, the weight of connectivity in the overall score rose from 10 to 20 % between 2012 and 2014, while resources and environment decreased by 5 % each. The position of Taiwan and Thailand improved dramatically in Connectivity from 2013 to 2014. To what extent can these changes be thanked to the changes of methods?
- the problem of distances: in many cases, the difference between overall results seems to be statistically insignificant. The difference between the score of Canada (3rd position in 2014) and the Netherlands (7th position) is 2.5 points on a scale of 100.
- Composite overall scores hide the differences between systems. Countries with different profiles are ranked similarly. For example, Finland and Denmark are very close to each other in the 2014 Ranking, but Finland has an advantage in Environment, while Denmark is much stronger in Connectivity and Resources (Table 1).

In my opinion, the impact of a system ranking is less direct as it has fewer direct consequences, so distortions caused by methodological problems are not as a dire problem as in the case of institutional rankings. No ministers will be relieved if a country falls back. The good brand and the high prestige of a system are hard to convert into monetary advantages.

Table 1 The change of position of selected countries between 2012 and 2014

	Ranking position in			Change between	
	2012	2013	2014	2012/2013	2013/2014
Norway	7	12	11	5	−1
Taiwan-China	21	26	22	5	−4
Spain	24	20	23	−4	3
Ukraine	25	36	43	11	7
Slovenia	28	23	25	−5	2
Bulgaria	31	38	40	7	2
Hungary	34	34	29	0	−5
Malaysia	36	27	28	−9	1
China	39	42	35	3	−7
Thailand	41	47	42	6	−5
Iran	42	48	49	6	1

That also means, however, that there are no real arguments for *ranking* systems, because rankings condense and therefore lose information. Thus, by providing only rankings, U21 fails to truly grab the diversity of higher education systems. Providing comparable indicators (rather than dimensions) and a *classification* of systems would be more useful and informative. U21 currently tells us which system is better (based on their calculation) but it provides only shallow information on why one system is considered better than the other.

This can be supported by Millot (2014), who argues that there is a strong correlation between the ranking of U21 and the density of institutions⁴ according to ARWU. The repetition of his calculation on ranking data of 2014 (see Appendix 1) shows the result of 0.91. This is a strong correlation, that is, based on ARWU and population numbers, U21 ranking is highly predictable. This is not surprising, if we take into consideration that U21 uses 7 indicators (weighting around 30 % altogether) to assess research in addition to the direct incorporation of ARWU results (2 indicators with a weight of 6.5 %).

5 U-Multirank

5.1 The General Characteristics and Strengths of U-Multirank

U-Multirank is designed and led by a consortium and its foundation was funded by the European Union. The consortium includes the Centre for Higher Education (CHE) in Germany, which runs rankings similar to U-Multirank in

⁴The number of institutions listed in ARWU 500 divided by the population of the country.

German-language countries, the Center for Higher Education Policy Studies (CHEPS) in the Netherlands, one of the most prominent European higher education research groups, the Centre for Science and Technology Studies from Leiden University, the producer of the Leiden Ranking, as well as other partners. Representatives of stakeholders (such as the European Student Union) were also involved. After a pilot phase, the first U-Multirank was officially published in 2014.

The U-Multirank ranks whole institutions as well as fields of education provided by institutions. Currently, business administration, mechanical engineering, electrical engineering physics are included, and in 2015 psychology, computer science and medicine will be involved.

U-Multirank differentiates itself from other rankings by defining itself as multi-dimensional and user-driven.

In U-Multirank 50 performance indicators in 5 dimensions (teaching and learning, research, knowledge transfer, international orientation and regional engagement) are defined. Being multi-dimensional means that U-Multirank does not create any composite score. There is no overall score for institutions, instead it is possible to create ranking for each of the performance indicators. The rationale behind this is that each user can select his/her own most important aspect making U-Multirank user-driven.

What is more, having selected an indicator, U-Multirank does not *rank* institutions, but *rate* them by grouping them into 5 broad categories, A to E, where A stands for very good, while E stands for weak. Some other rankings also categorize institutions, but the categorization is usually based on the ranking position itself. Here, in most cases categorization is based on to what extent a score of an institution differs from the group median. There is no ranking within the categories (institutions appear in alphabetical order).

That approach makes U-Multirank less hierarchical and less competitive, because half of the institutions with non-zero score will achieve A or B, and a lot of institutions can be rated A at the same time.⁵ It also makes U-Multirank less sensitive for errors stemming from insignificant statistical differences and also for changing methodology, as new indicators can be introduced without disturbing the existing ones.

The less competitive nature of U-Multirank makes it possible for institutions to follow their own strategy. It is not the rankings which form the profile and strategy of institutions any more, because institutions do not need to be good in *all* indicators to be ranked well (as it is the case with traditional rankings). Institutions can select only those indicators which fit their own strategy. In a presentation, Frans van Vught⁶ emphasised that U-Multirank made several institutions visible, who performed excellent in one or more indicators and who are not able to compete in

⁵A competition for collecting as many as possible may emerge. However, the more performance indicators are introduced, the more difficult to compete in each “event”. In addition, there are indicators which mutually excludes each other (such as number of international students and the BA graduates working in region).

⁶Presentation held at the University of Twente, 15 May 2014.

traditional, one-dimensional and hierarchical rankings. Thus, U-Multirank truly gives space for diversity, and does not force institutions explicitly or implicitly to follow one predefined script.

This is enhanced even more by creating mapping indicators in addition to performance indicators (such as size, age, income from different sources, broad subject areas, etc.), which are used to describe the activity profile of each institution. That makes it possible to compare only those institutions which are similar.

5.2 Challenges

Several indicators defined by U-Multirank are rarely used in other rankings. For example, graduating on time, number of spin-offs or student mobility are hardly ever seen in the most popular global rankings. In general, U-Multirank indicators cover third mission activities much more than any other rankings.

Some of the newer indicators, however, depend only partially on institutions, and policy context has much influence on them. For example “graduation on time” depends on admission and selectivity rules. Thus, U-Multirank is not sensitive for different policy contexts, consequently the ranking could be improved further by including mapping indicators regarding the system level.

To produce indicator, U-Multirank collects data

- from existing databases, such as international publication database (Web of Science) or patent database (EPO Worldwide Patent Statistical Database),
- from students by student satisfaction surveys and
- from institutions through institutional self-reporting (institutional and field-based questionnaires).

Rauhvargers (2013) criticizes U-Multirank that it chose Web of Science alone. Scopus covers more journals and types of publications, which would have fit better in the more inclusive approach of U-Multirank.

Although U-Multirank does not use reputational surveys, indicators regarding the environment of teaching and learning are based on surveys among students. Comparison of the results of such surveys internationally or even among institutions is questionable because responses are based on previous *expectations* which are influenced by many factors. For example, the high prestige of an institution might create false expectations. Such an institution might do poorer in the survey than the less prestigious ones, even if the quality of the institution is better from an objective perspective.⁷

⁷For instance, the indicator of “quality of courses and teaching” includes responses to the question of “option to choose elective courses”. With low level of expectations, students might be highly satisfied with few options, while high level of expectation results in low satisfaction even if the number of courses is higher. This is also true for the other components of the indicator, that is, for “the breadth of teaching offerings”, “the didactic quality of teaching”, “the quality of basic

Nevertheless, current students' satisfaction could be helpful for prospective students even if direct comparison is problematic. The length of the questionnaire (more than 100 items) and its limited language availability (it is translated only into English, French, German, Spanish, Polish, and Russian) might cause further distortions and the exclusion of less internationalized institutions and disciplines. It is worth noting, however, that the survey has a very low break up rate.⁸

In addition to student surveys and international databases, the majority of indicators require institutional self-reporting. In a case as large as U-Multirank, this can be a serious problem, not particularly because of data manipulation, but because of the lack of consistency, especially in the case of regional engagement and knowledge transfer indicators. Achieving a common understanding of "private sources" for calculating "Income from private sources" or "region" for "BA graduates working in region", for example, requires a lot of discussions. Even producing the raw data for some indicators can be a challenge for many institutions (e.g. Art related output). On the other hand, however, by defining new but relevant indicators, U-Multirank "educates" institutions and helps them to institutionalize data gathering processes and makes them capable of revealing less known aspects of their activity for the public.

The mission of U-Multirank is to provide such a transparency tool that does not constrain institutional diversity, but promotes benchmarking and competition. This mission can be fulfilled only if as many institutions participate in the ranking, as possible. Increasing the number of participants (and indicators), however, also increases difficulties in maintaining consistency. Therefore U-Multirank is an attempt to surpass what Stella and Woodhouse think hopeless: "since rankings also imply that the whole system has to be covered within a time frame, it would be futile to attempt in a large and complex system. At the most, it can be done only at a superficial level, akin to the methodology followed by the media. Consequently, lack of validation of self-reported data, inconsistency in terminologies, lack of peer review, inability to consider institutional diversities, etc. would become unavoidable, thus rendering the outcome of the whole process useless." (Harvey 2008).

Current practices followed by U-Multirank regarding consistency are less transparent. U-Multirank describes this process in the following way: "To ensure comparability of data across institutions, the questionnaires include guidelines and definitions of all data items requested. [...] Data are then intensively checked by the U-Multirank team, applying both automated and manual checks for consistency, plausibility (including checks of outliers) and missing data."⁹ This is followed by an

(Footnote 7 continued)

courses", etc. In sum, the indicator of "quality of courses and teaching" does not reflect on the quality per se (i.e. the number of optional courses), but on to what extent the institution meets students' expectations regarding teaching. We know nothing about these expectations, however, which makes comparison of institutions based on the "quality of courses and teaching" indicator dubious.

⁸Gero Federkeil, written communication.

⁹<http://www.umultirank.org/#!/methodology?section=undefined> (accessed 23 Sept 2014).

iterative process between U-Multirank team and institutional representatives aiming to clarify and to correct data.

It is obvious that the higher the number of participants is, the more resources are required to maintain consistency. Therefore financial sustainability of U-Multirank is an important question to be raised. The U-Multirank consortium explored several governance and funding options in the feasibility study (van Vught and Ziegele 2011), and they support the idea of an independent, non-profit provider funded by European Commission and other foundations with different market sources (extra services, advertisement, subscription fees).

Costs of maintaining consistency could be decreased by involving national statistical agencies. Taking into consideration the depth of required data, this could be carried out primarily on European level, the possibility of which was examined in another EU-project (called EUMIDA). A potential risk could be the isolation of non-European institutions. Charging fees for participation can also decrease the motivation of institutions to provide data. Additional income could come from extra services, such as data clearing activities, when U-Multirank collects special data from institutions and then reports them back for non-public benchmarking purposes.

The number of active participants determines the success of U-Multirank. Participation depends on what costs and benefits U-Multirank causes to institutions. On one hand, self-reporting generates high workload for institutions. On the other hand, U-Multirank provides some possibilities for benchmarking, which could be enhanced even further, if U-Multirank provides access to more personalised and more detailed comparative data for participating institutions. (That could be an additional source of income.)

From the institutional point of view, an additional benefit could be the possibility to increase recruitment and mobility. The first round of U-Multirank is quite European-focused. Although U-Multirank emphasises that the number of ranked institutions is more than 850, the number of active participants who actually provided data is much lower, it is around 500. Data for the rest of the institutions comes from international (mainly bibliographic) databases. The majority of active participants are located in the EU (382 institutions) or in the broader region of the European Higher Education Area (48 institutions). Only 74 institutions can be found in the other parts of the world. Some countries are significantly underrepresented in the ranking: there are only 9-9 institutions from the United States and UK respectively, China is represented by 4 institutions and Canada by 2. (For further details and the calculation of numbers, see Appendix 2) For the non-European institutions, the benefit of increased recruitment is viable only if U-Multirank becomes truly global, and a critical mass of non-European institutions is reached.

Those institutions that have no chance to appear near the top of the current global rankings (or on the rankings at all), might be interested in being present in U-Multirank, which is a more democratic ranking than the traditional global ones. On the other hand, it also generates less prestige for participants because there are a lot of winners. It also undermines the position of universities heading the current global

rankings. Therefore, neither the hostile reactions from League of European Research Universities (LERU) and Russell Group (consisting of UK-based research intensive universities), nor the absence of US universities among the active participants is surprising. In the long run, however, winning research intensive universities over, in order to participate in U-Multirank is inevitable. No rankings without them would be credible. Reaching critical mass might help to convince them.

The multidimensional way that U-Multirank follows has its potential risks. It is a question whether users, students are prepared enough for a ranking with several winners or they rather continue looking for “the best”. U-Multirank requires users to have clear priorities and a certain level of maturity. Without that, the potential danger of crowding out effect might occur, that is, a simple ranking which requires less effort from the user crowds out the more sophisticated, better rankings which require more efforts. This can occur if users have no information about the quality of different rankings, thus he/she cannot make a difference between them. Informing and educating the public is therefore crucial for the success of U-Multirank.

Current rankings can easily copy some features of U-Multirank. The strengths of U-Multirank lies in its unique database, in the classification system, its approach of rating rather than ranking, and the interactive, user driven service which makes the creation of more personalized rankings possible. Providing field-based ranking in addition to institutional rankings is also an important, but less unique characteristic.

With the exception of the unique database, much of these can be copied or imitated obscuring the real differences. While maintaining their existing, authoritative “overall rankings”, global ranking providers can create more interactive services, where users can set their own weightings, can rank according to specific indicators, etc. (Vercruysse and Proteasa 2012). Some of the providers have already started to develop simple classification systems. Rating can be introduced without losing much competitiveness provided high number of categories is defined. Field based rankings are easily replaced by currently existing subject rankings, etc.

It is worth noting, however, that if traditional ranking providers introduce all the changes above, it will change the market of rankings considerably. However, from the U-Multirank’s point of view, the question is whether better data and more relevant indicators are important enough for the mass of users and institutions, so that maintaining U-Multirank would be worth in the future.

6 Conclusions: Rankings and the European Higher Education Area

The European Higher Education has some distinctive characteristics which made the penetration of rankings more difficult. In the US rankings spread in the 1980s, while in Europe it was the period of 1990s and 2000s when most national rankings appeared. Why?

One possible explanation could be that European higher education systems are less competitive, less hierarchical and more egalitarian. As the state plays a major role in maintaining institutions, quality differences between them are less tolerated and less obvious. In some countries there are predefined categories for institutions (i.e. universities and colleges), but within categories, there is less possibility (authority, resources) for institutions to make real differences.

Interestingly, national rankings rarely had an impact on policy making. The situation changed with the appearance of global rankings, particularly ARWU and THE-Thomson Reuters World University Rankings proved to be influential on the European Higher Education Area (Vercruysse and Proteasa 2012). Although their results are mostly distorted because of the ignorance of methodological concerns, they raised serious questions regarding the competitiveness of European higher education institutions, and encouraged governments to intervene.¹⁰ These interventions, aiming to make world class universities, create tension within higher education systems.

Another important characteristic is the simultaneous presence of convergence and divergence. The former stems from such European initiatives like the Bologna process, while the latter is because of path dependency. Policy practices work differently in different contexts, and European countries have their own history and identity. That makes European higher education more complex and more diverse on one hand, and less transparent on the other.

It is imperative therefore to create more transparency to make differences in contexts as well as in performances more visible. The evidence-based approach of rankings is also necessary to provoke profound discussions on higher education systems and institutions. Another goal of creating transparency is to promote and provoke competition, but that should be done without enforcing predefined scripts. Institutions must have the possibility to choose their strategies and rankings (as well as other transparency tools) and it must be respected. Transparency should be supported by developing incentive structures which reward all aspects of institutional performance (including teaching and third mission activities) (Hazelkorn 2012).

U21 and U-Multirank are promising new tools wishing to meet these aims, even if they have their own challenges and weaknesses. Their success depends on whether the public understands and appreciates key differences between them and the more traditional global rankings.

Acknowledgements I wish to express my sincere gratitude to Jan Sadlak and Gero Federkeil for their comments and suggestions. I also thank to Ms. Zsuzsanna Döme for her help in translating the paper.

Open Access This chapter is distributed under the terms of the Creative Commons Attribution Noncommercial License, which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

¹⁰As Vercruysse and Proteasa (2012) mention, it is interesting that national rankings rarely have an impact on policy making. International (global) rankings however seem to be more influential.

Appendix 1

Correlation between U21 and density of institutions according to ARWU 2014

	Rank in U21 2014	Number of institutions in ARWU 2014 (TOP 500)	Population in million (2013) ^a	Density of ARWU institutions (ARWU institution/population)	Rank according to ARWU density
Argentina	41	1	41.4	0.02	36
Australia	9	19	23.3	0.82	6
Austria	12	6	8.5	0.71	9
Belgium	13	7	11.1	0.63	11
Brazil	38	6	200.4	0.03	35
Bulgaria	40	0	7.2	0.00	43
Canada	3	21	35.2	0.60	14
Chile	33	2	17.6	0.11	28
China	35	44	1385.6	0.03	34
Croatia	44	0	4.3	0.00	47
Czech Republic	26	1	10.7	0.09	30
Denmark	4	5	5.6	0.89	3
Finland	5	5	5.4	0.93	2
France	18	21	64.3	0.33	20
Germany	14	39	82.7	0.47	16
Greece	32	2	11.1	0.18	25
Hungary	29	2	10	0.20	24
India	50	1	1252.1	0.00	41
Indonesia	48	0	249.9	0.00	48
Iran	49	1	77.4	0.01	39
Ireland	17	3	4.6	0.65	10
Israel	19	6	7.7	0.78	7
Italy	27	21	61	0.34	19
Japan	20	19	127.1	0.15	26
Korea	21	10	49.3	0.20	23
Malaysia	28	2	29.7	0.07	32
Mexico	46	1	122.3	0.01	40
Netherlands	7	13	16.8	0.77	8
New Zealand	16	4	4.5	0.89	4
Norway	11	3	5	0.60	13
Poland	31	2	38.2	0.05	33
Portugal	24	3	10.6	0.28	21
Romania	39	0	21.7	0.00	44
Russian Federation	36	2	142.8	0.01	37

(continued)

(continued)

	Rank in U21 2014	Number of institutions in ARWU 2014 (TOP 500)	Population in million (2013) ^a	Density of ARWU institutions (ARWU institution/population)	Rank according to ARWU density
Saudi Arabia	30	4	28.8	0.14	27
Serbia	34	1	9.5	0.11	29
Singapore	10	2	5.4	0.37	18
Slovakia	37	0	5.5	0.00	45
Slovenia	25	1	2.1	0.48	15
South Africa	45	4	52.8	0.08	31
Spain	23	12	46.9	0.26	22
Sweden	2	11	9.6	1.15	1
Switzerland	6	7	8.1	0.86	5
Thailand	42	0	67	0.00	46
Turkey	47	1	74.9	0.01	38
Ukraine	43	0	45.2	0.00	42
United Kingdom	8	38	63.1	0.60	12
USA	1	146	320.1	0.46	17

^aSource http://en.wikipedia.org/wiki/List_of_countries_by_population_%28United_Nations%29

For the sake of convenience, I left Taiwan and Hong Kong SAR out of calculation (there are no separate ARWU data regarding these countries)

Correlation of Rank in U21 (2014) and rank according to density of ARWU 2014 is 0.909169

Appendix 2

Currently, there are no country statistics available on umultirank.org. The number of active participants is calculated by counting institutions offering bachelor and master degrees, with relevant data in one of the following indicators: Bachelor graduation rate, Masters graduation rate, Graduating on time (bachelors), Graduating on time (masters), Income from private sources, Spin-offs, Student mobility, Bachelor graduates working in the region, Master graduates working in the region, Student internships in the region, Income from regional sources.

Number of active participants by country

Country	Number of institutions	Country	Number of institutions
Austria	13	Canada	2
Belgium	6	USA	9
Bulgaria	5	North-America	11

(continued)

(continued)

Country	Number of institutions	Country	Number of institutions
Croatia	2	Brazil	3
Cyprus	5	Chile	3
Czech Republic	14	Costa Rica	1
Denmark	9	Jamaica	1
Estonia	2	Mexico	2
Finland	16	Peru	1
France	57	Uruguay	1
Germany	37	South-America	12
Greece	9	Iran	1
Hungary	10	Israel	1
Ireland	8	Palestine	1
Italy	23	Saudi Arabia	2
Latvia	10	United Arab Emirates	1
Liechtenstein	1	Close-East	6
Lithuania	9	Bangladesh	1
Malta	1	China	4
Netherlands	15	Chinese Taipei	1
Poland	34	India	3
Portugal	20	Japan	6
Romania	10	Malaysia	1
Slovakia	6	South Korea	1
Slovenia	3	Thailand	1
Spain	39	Vietnam	1
Sweden	9	Far-East	19
United Kingdom	9	Australia	11
EU-members	382	New Zealand	1
Iceland	1	Oceania	12
Kazakhstan	3	Ghana	1
Macedonia (fYRoM)	2	Morocco	1
Norway	11	Namibia	1
Russia	9	Nigeria	2
Serbia	2	South Africa	6
Switzerland	6	Africa	11
Turkey	7	Azerbajdjan	1
Ukraine	7	Belorus	2
EHEA members without EU membership	48	Other countries	3

Source umultirank.org (accessed: 24 Sept 2014)

References

- Aula, H.-M., & Tienari, J. (2011). Becoming “world-class”? Reputation-building in a university merger. *Critical Perspectives on International Business*, 7(1), 7–29.
- Deephouse, D. L., & Suchman, M. (2006). Legitimacy in organizational institutionalism. In R. Greenwood, C. Oliver, R. Suddaby, & K. Sahlin (Eds.), *The SAGE handbook of organizational institutionalism* (pp. 49–77). London: Sage Publications.
- Eckel, P. D. (2008). Mission diversity and the tension between prestige and effectiveness: An overview of US higher education. *Higher Education Policy*, 21, 175–192.
- Gibbons, M., Limoges, L., & Nowotny, H. (1994). *The new production of knowledge. The dynamics of science and research in contemporary societies*. London: Sage Publication.
- Harvey, L. (2008). Ranking of higher education institutions. In A. Stella & D. Woodhouse (Eds.), *Occasional Publications Series no. 6* (2006 ed.). Melbourne: AUQA.
- Hazelkorn, E. (2011). *Rankings and the reshaping of higher education. The Battle for world-class excellence*. London: Palgrave-McMillan.
- Hazelkorn, E. (2012). European “transparency instruments”: Driving the modernisation of European higher education. In Adrian Curaj et al. (Ed.), *European higher education at the crossroads: Between the Bologna Process and National Reforms* (pp. 339–360). Heidelberg New York London: Dordrecht.
- Millot, B. (2014). Rankings—Higher education systems vs. universities. *University world news*. 312. Retrieved September 25 2014, from <http://www.universityworldnews.com/article.php?story=20140318125156290>.
- Rauhvargers, A. (2011). *Global university rankings and their impact*. Retrieved from http://www.eua.be/pubs/global_university_rankings_and_their_impact.pdf.
- Rauhvargers, A. (2013). *Global university rankings and their impact. Report II*. Retrieved from http://www.eua.be/Libraries/Publications_homepage_list/EUA_Global_University_Rankings_and_Their_Impact_-_Report_II.sfb.ashx.
- Sadlak, J. (2014). University rankings: The manifestation and driver of competition for excellence within the new higher education landscape. In P. Mattei (Ed.), *University adaptation in difficult economic times* (pp. 137–154). Oxford/New York: Oxford University Press.
- Salmi, J., & Saroyan, A. (2007). League tables as policy instruments: Uses and misuses. *Higher Education Management and Policy*, 19(2), 31–68.
- Soh, K. C. (2011). Don’t read university rankings like reading football league tables: Taking a close look at the indicators. *Higher Education Reviews*, 44(1), 15–29.
- Soh, K. C. (2012). The U21 rankings of national higher education systems: Re-analyze to optimize. *European Journal of Higher Education*, 2(1–2), 205–217.
- van der Wende, M. (2008). Rankings and classifications in higher education: A European perspective. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 23, pp. 49–71).
- van Vught, F. (2008). Mission diversity and reputation in higher education. *Higher Education Policy*, 21, 151–174.
- van Vught, F., & Ziegele, F. (2011). *Design and testing the feasibility of a multidimensional global university ranking. Final report*. Retrieved September 25, 2014, from ec.europa.eu/education/library/study/2011/multirank_en.pdf.
- Vercruyssen, N., & Proteasa, V. (2012). *Transparency tools across the European higher education area*. Retrieved from http://www.ehea.info/Uploads/%281%29/brosura_v1_v12_vp_120419_text.pdf.
- Williams, R., de Rassenfossé, G., Jensen, P., Marginson, S. (2014). *U21 ranking of national higher education systems*. Melbourne Institute. Available at: <http://www.universitas21.com/article/projects/details/153/executive-summary-and-full-2013-report>, www.umultirank.org.